5″

TTGACACCAG	ACCAACTGGT	AATGGTAGCG	ACCGGCGCTC	AGCTGGAATT	CCAAAAAATG
TAATGCACAC	TCCATTGCAT	TCAGCCCGCC	TCTCCTTAGT	CGCCGCCATG	ACGACCGCGT
CCACCTCGCA	GGTGCGCCAG	AACTACCACC	AGGACTCAGA	GGCCGCCATC	AACCGCCAGA
TCAACCTGGA	GCTCTACGCC	TCCTACGTTT	ACCTGTCCAT	GTCTTACTAC	TTTGACCGCG
ATGATGTGGC	TTTGAAGAAC	TTTGCCAAAT	ACTTTCTTCA	CCAATCTCAT	GAGGAGAGGG
AACATGCTGA	GAAACTGATG	AAGCTGCAGA	ACCAACGAGG	TGGCCGAATC	TTCCTTCAGG
ATATCAAGAA	ACCAGACTGT	GATGACTGGG	AGAGCGGGCT	GAATGCAATG	GAGTGTGCAT
TACATTTGGA	AAAAAATGTG	AATCAGTCAC	TACTGGAATT	CCCTTCTCCT	ATCTCTCCCA
GTCCTAGCTG	CTGGCATCAC	TATACTACTA	ACAGACCGCA	ACCTCAACAC	CACCTTCTTC
GACCCCGCCG	GAGGAAGAGA	CCCCATTCTA	TACCAACACC	TATTCTGATT	TTTCGGTCAC
CCTGAAGTTT	ATATTCTTAT	CCTACCAGGC	TTCGGAATAA	TCTCCCATAT	TGTAACTTAC
TACTCCGGAA	ATCGCTGTCG	CCTAACCGCT	AACATTACTG	CAGGCCACCT	ACTCATGCAC
CTAATTGGAA	GCGCCACCCT	AGCAATATCA	ACCATTAACC	TTCCCTCTAC	ACTTATCATC
TTCACAATTC	TAATTCTACT	GACTATCCTA	GAAATCGCTG	TCGCCTTAAT	CCAAGCCTAC
GTTTTCACAC	TTCTAGTAAG	CCTCTACCTG	CACGACAACA	CATAAAAAAA	A 3″

FIG. 1

CLONE p47

GGGGGACGGAACCCGG

CGCTCGTTCCCCACCCCGGCCGGCCGCCCATAGCCAGCCCTCCGTCAC

CLONE T 16

TTGACACC

CTCTTCACCGCACCCTCGGACTGCCCCAAGGCCCCCGCCGCCGCTCC

AGACCAACTGGTAATGGTAGCGACCGGCGCTCAGCTGGAATTCCAAAA

AGCGCCGCGCAGCCACCGCCGCCGCCGCCGCCTCTCCTTAGTCGCCGCC

AATGTAATGCACACTCCATTGCATTCAGCCCGCCTCTCCTTAGTCGCCGCC

ATG	ACG	ACC	GCG	TCC	ACC	TCG	CAG	GTG	CGC	CAG
ATG	ACG	ACC	GCG	TCC	ACC	TCG	CAG	GTG	CGC	CAG
AAC	TAC	CAC	CAG	GAC	TCA	GAG	GCC	GCC	ATC	AAC
AAC	TAC	CAC	CAG	GAC	TCA	GAG	GCC	GCC	ATC	AAC
CGC	CAG	ATC	AAC	CTG	GAG	CTC	TAC	GCC	TCC	TAC
CGC	CAG	ATC	AAC	CTG	GAG	CTC	TAC	GCC	TCC	TAC
GTT	TAC	CTG	TCC	ATG	TCT	TAC	TAC	TTT	GAC	CGC
GTT	TAC	CTG	TCC	ATG	TCT	TAC	TAC	TTT	GAC	CGC
GAT	GAT	GTG	GCT	TTG	AAG	AAC	TTT	GCC	AAA	TAC
GAT	GAT	GTG	GCT	TTG	AAG	AAC	TTT	GCC	AAA	TAC
TTT	CTT	CAC	CAA	TCT	CAT	GAG	GAG	AGG	GAA	CAT
TTT	CTT	CAC	CAA	TCT	CAT	GAG	GAG	AGG	GAA	CAT
GCT	GAG	AAA	CTG	ATG	AAG	CTG	CAG	AAC	CAA	CGA
GCT	GAG	AAA	CTG	ATG	AAG	CTG	CAG	AAC	CAA	CGA
GGT	GGC	CGA	ATC	TTC	CTT	CAG	GAT	ATC	AAG	AAA
GGT	GGC	CGA	ATC	TTC	CTT	CAG	GAT	ATC	AAG	AAA
CCA	GAC	TGT	GAT	GAC	TGG	GAG	AGC	GGG	CTG	AAT
CCA	GAC	TGT	GAT	GAC	TGG	GAG	AGC	GGG	CTG	AAT
GCA	ATG	GAG	TGT	GCA	TTA	CAT	TTG	GAA	AAA	AAT
GCA	ATG	GAG	TGT	GCA	TTA	CAT	TTG	GAA	AAA	AAT
GTG	AAT	CAG	TCA	CTA	CTG	GAA	CTG	CAC	AAA	CTG
GTG	AAT	CAG	TCA	CTA	CTG	GAA	TTC	CCT	TCT	CCT
GCC	ACT	GAC	AAA	AAT	GAC	CCC	CAT	TTG	TGT	GAC
ATC	TCT	CCC	AGT	CCT	AGC	TGC	TGG	CAT	CAC	TAT

FIG. 2A

TTC ATT GAG ACA CAT TAC CTG AAT GAG CAG GTG
ACT ACT AAC AGA CCG CAA CCT CAA CAC CAC CTT
AAA GCC ATC AAA GAA TTG GGT GAC CAC GTG ACC
CTT CGA CCC CGC CGG AGG AAG AGA CCC CAT TCT
AAC TTG CGC AAG ATG GGA GCG CCC GAA TCT GGC
ATA CCA ACA CCT ATT CTG ATT TTT CGG TCA CCC
TTG GCG GAA TAT CTC TTT GAC AAG CAC ACC CTG
TGA AGTTTATATTCTTATCCTACCAGGCTTCGGAATAATCTCCCATATT

GGA GAC AGT GAT AAT GAA AGC TAA GCCTCGGGCTAATT
GTAACTTACTACTCCGGAAATCGCTGTCGCCTAACCGCTAACATTACTGC

TCCCATAGCCGTGGGGTGACTTCCCTGGTCACCAAGGCAGTGCATGCAT
AGGCCACCTACTCATGCACCTAATTGGAAGCGCCACCCTAGCAATATCA

GCATGTTGGGGTTTCCTTTACCTTTTCTATAAGTTGTACCAAAACATCCAC
ACCATTAACCTTCCCTCTACACTTATCATCTTCACAATTCTAATTCTACTG

TTAAGTTCTTTGATTTGTACCATTCCTTCAAATAAAGAAATTTGGTACCCA
ACTATCCTAGAAATCGCTGTCGCCTTAATCCAAGCCTACGTTTTCACACT

AAAAAAAA
TCTAGTAAGCCTCTACCTGCACGACAACACATAAAAAAAA

FIG. 2A CONT.

TTGACACCAGACCAACTGGTAATGGTAGCGACCGGCGCTCAGCTGGAATTCCAAAAAATGT
AATGCACACTCCATTGCATTCAGCCCGCCTCTCCTTAGTCGCCGCC

met	thr	thr	ala	ser	thr	ser	gln	val	arg	gln
ATG	ACG	ACC	GCG	TCC	ACC	TCG	CAG	GTG	CGC	CAG
asn	tyr	his	gln	asp	ser	glu	ala	ala	ile	asn
AAC	TAC	CAC	CAG	GAC	TCA	GAG	GCC	GCC	ATC	AAC
arg	gln	ile	asn	leu	glu	leu	tyr	ala	ser	tyr
CGC	CAG	ATC	AAC	CTG	GAG	CTC	TAC	GCC	TCC	TAC
val	tyr	leu	ser	met	ser	tyr	tyr	phe	asp	arg
GTT	TAC	CTG	TCC	ATG	TCT	TAC	TAC	TTT	GAC	CGC
asp	asp	val	ala	leu	lys	asn	phe	ala	lys	tyr
GAT	GAT	GTG	GCT	TTG	AAG	AAC	TTT	GCC	AAA	TAC
phe	leu	his	gln	ser	his	glu	glu	arg	glu	his
TTT	CTT	CAC	CAA	TCT	CAT	GAG	GAG	AGG	GAA	CAT
ala	glu	lys	leu	met	lys	leu	gln	asn	gln	arg
GCT	GAG	AAA	CTG	ATG	AAG	CTG	CAG	AAC	CAA	CGA
gly	gly	arg	ile	phe	leu	gln	asp	ile	lys	lys
GGT	GGC	CGA	ATC	TTC	CTT	CAG	GAT	ATC	AAG	AAA
pro	asp	cys	asp	asp	trp	glu	ser	gly	leu	asn
CCA	GAC	TGT	GAT	GAC	TGG	GAG	AGC	GGG	CTG	AAT
ala	met	glu	cys	ala	leu	his	leu	glu	lys	asn
GCA	ATG	GAG	TGT	GCA	TTA	CAT	TTG	GAA	AAA	AAT
val	asn	gln	ser	leu	leu	glu	phe	pro	ser	pro
GTG	AAT	CAG	TCA	CTA	CTG	GAA	TTC	CCT	TCT	CCT
ile	ser	pro	ser	pro	ser	cys	trp	his	his	tyr
ATC	TCT	CCC	AGT	CCT	AGC	TGC	TGG	CAT	CAC	TAT
thr	thr	asn	arg	pro	gln	pro	gln	his	his	leu
ACT	ACT	AAC	AGA	CCG	CAA	CCT	CAA	CAC	CAC	CTT
leu	arg	pro	arg	arg	arg	lys	arg	pro	his	ser
CTT	CGA	CCC	CGC	CGG	AGG	AAG	AGA	CCC	CAT	TCT
ile	pro	thr	pro	ile	leu	ile	phe	arg	ser	pro
ATA	CCA	ACA	CCT	ATT	CTG	ATT	TTT	CGG	TCA	CCC

TGA AGTTTATATTCTTATCCTACCAGGCTTCGGAATAATCTCCCATATTGTAACTTAC
TACTCCGGAAATCGCTGTCGCCTAACCGCTAACATTACTGCAGGCCACCTACTCATGCAC
CTAATTGGAAGCGCCACCCTAGCAATATCAACCATTAACCTTCCCTCTACACTTATCATC
TTCACAATTCTAATTCTACTGACTATCCTAGAAATCGCTGTCGCCTTAATCCAAGCCTAC
GTTTTCACACTTGTAGTAAGCCTCTACCTGCACGACAACACATAAAAAAAA

FIG. 5

1061 BNC

TTGACACCAG ACCAACTGGT AATGGTAGCG ACCGGCGCTC AGCTGGAATT CCAAAAAAATG

NCS

TAATGCACAC TCCATTGCAT TCAGCCCGCC TCTCCTTAGT CGCCGCCATG ACGACCGCGT

X1

CCACCTCGCA GGTGCGCCAG AACTACCACC AGGACTCAGA GGCCGCCATG AACCGCCAGA

17

TCAACCTGGA GCTCTACGCC TCCTACGTTT ACCTGTCCAT GTCTTACTAC TTTGACCGCG

17

ATGATGTGGC TTTGAAGAAC TTTGCCAAAT ACTTTCTTCA CCAATCTCAT GAGGAGAGGG

AACATGCTGA GAAACTGATG AAGCTGCAGA ACCAACGAGG TGGCCGAATC TTCCTTCAGG

2.1

ATATCAAGAA ACCAGACTGT GATGACTGGG AGAGCGGGCT GAATGCAATG GAGTGTGCAT

ECOF

TACATTTGGA AAAAAATGTG AATCAGTCAC TACTGGAATT CCCTTCTCCT ATCTCTCCCA

GTCCTAGCTG CTGGCATCAC TATACTACTA ACAGACCGCA ACCTCAACAC CACCTTCTTC

GACCCCGCCG GAGGAAGAGA CCCCATTCTA TACCAACACC TATTCTGATT TTTCGGTCAC

CCTGAAGTTT ATATTCTTAT CCTACCAGGC TTCGGAATAA TCTCCCATAT TGTAACTTAC

SPF

TACTCCGGAA ATCGCTGTCG CCTAACCGCT AACATTACTG CAGGCCACCT ACTCATGCAC

728 767

CTAATTGGAA GCGCCACCCT AGCAATATCA ACCATTAACC TTCCCTCTAC ACTTATCATC

767 16

TTCACAATTC TAATTCTACT GACTATCCTA GAAATCGCTG TCGCCTTAAT CCAAGCCTAC

GTTTTCACAC TTCTAGTAAG CCTCTACCTG CACGACAACA CATAAAAAAA A

FIG. 7

TTGACACCAG ACCAACTGGT AATGGTAGCG ACCGGCGCTC AGCTGGAATT CCAAAAAATG

TAATGCACAC TCCATTGCAT TCAGCCCGCC TCTCCTTAGT CGCCGCCATG ACGACCGCGT

CCACCTCGCA GGTGCGCCAG AACTACCACC AGGACTCAGA GGCCGCCATC AACCGCCAGA

TCAACCTGGA GCTCTACGCC TCCTACGTTT ACCTGTCCAT GTCTTACTAC TTTGACCGCG

ATGATGTGGC TTTGAAGAAC TTTGCCAAAT ACTTTCTTCA CCAATCTCAT GAGGAGAGGG

Pst1
AACATGCTGA GAAACTGATG AAGCTGCAGA ACCAACGAGG TGGCCGAATC TTCCTTCAGG

ATATCAAGAA ACCAGACTGT GATGACTGGG AGAGCGGGCT GAATGCAATG GAGTGTGCAT

ECOR1
TACATTTGGA AAAAAATGTG AATCAGTCAC TACTGGAATT CCCTTCTCCT ATCTCTCCCA

GTCCTAGCTG CTGGCATCAC TATACTACTA ACAGACCGCA ACCTCAACAC CACCTTCTTC

GACCCCGCCG GAGGAAGAGA CCCCATTCTA TACCAACACC TATTCTGATT TTTCGGTCAC

CCTGAAGTTT ATATTCTTAT CCTACCAGGC TTCGGAATAA TCTCCCATAT TGTAACTTAC

TACTCCGGAA ATCGCTGTCG CCTAACCGCT AACATTACTG CAGGCCACCT ACTCATGCAC

CTAATTGGAA GCGCCACCCT AGCAATATCA ACCATTAACC TTCCCTCTAC ACTTATCATC

TTCACAATTC TAATTCTACT GACTATCCTA GAAATCGCTG TCGCCTTAAT CCAAGCCTAC

GTTTTCACAC TTCTAGTAAG CCTCTACCTG CACGACAACA CATAAAAAAA A

FIG. 8